



Cornell University
ILR School

Cornell University ILR School
DigitalCommons@ILR

Articles and Chapters

ILR Collection

1979

Decomposing LDC Inequality

Gary S. Fields

Cornell University, gsf2@cornell.edu

Follow this and additional works at: <https://digitalcommons.ilr.cornell.edu/articles>



Part of the [Growth and Development Commons](#), [Income Distribution Commons](#), [International and Comparative Labor Relations Commons](#), and the [Labor Economics Commons](#)

Thank you for downloading an article from DigitalCommons@ILR.

Support this valuable resource today!

This Article is brought to you for free and open access by the ILR Collection at DigitalCommons@ILR. It has been accepted for inclusion in Articles and Chapters by an authorized administrator of DigitalCommons@ILR. For more information, please contact catherwood-dig@cornell.edu.

If you have a disability and are having trouble accessing information on this website or need materials in an alternate format, contact web-accessibility@cornell.edu for assistance.

Decomposing LDC Inequality

Abstract

[Excerpt] At the present time, there is great interest among development economists in the problem of economic inequality in less developed countries (LDCs). Studies of the determinants of inequality follow either of two general approaches. The more traditional approach is associated with names like Kuznets (1963), Chenery and associates (1960, 1968, 1975), Adelman and Morris (1973), Ahluwalia (1976) and Chiswick (1971). These studies share a common methodology, consisting basically of looking at a *cross-section of countries*, and (1) measuring the degree of inequality in each, (2) measuring other characteristics of each country (e.g., level of GNP, its rate of growth, importance of agriculture in total product, etc.), and (3) relating the level of inequality to that economy's characteristics using correlation or regression analysis.

In the last few years another type of approach has been followed, which looks instead at inequality *within a country*, and measures the contribution of the various components to total inequality. In this type of approach, using a variety of methodologies, inequality has been decomposed by economic sector (e.g., urban vs. rural), income source (e.g., income from labor vs. capital vs. land vs. transfers), or family characteristics (including attributes of the workers, their jobs, and regional and other locational considerations). This mode of inquiry is potentially of great value for understanding the structure of inequality and identifying which are the most important explanatory factors.

This study explores the decomposition type of inequality analysis. I summarize the alternative decomposition methodologies which have been set forth in the literature and review the principal findings of empirical studies.

Keywords

less developed countries, LDC, inequality, development

Disciplines

Growth and Development | Income Distribution | International and Comparative Labor Relations | Labor Economics

Comments

Required Publisher Statement

© [Oxford University Press](#). Final version published as: Fields, G. S. (1979). Decomposing LDC inequality. *Oxford Economic Papers*, 31(3), 437-459. Reprinted with permission. All rights reserved.

Suggested Citation

Fields, G. S. (1979). *Decomposing LDC inequality*[Electronic version]. Retrieved [insert date], from Cornell University, ILR School site:
<http://digitalcommons.ilr.cornell.edu/articles/978>

DECOMPOSING LDC INEQUALITY

By GARY S. FIELDS

I. Introduction

AT THE present time, there is great interest among development economists in the problem of economic inequality in less developed countries (LDCs). Studies of the determinants of inequality follow either of two general approaches. The more traditional approach is associated with names like Kuznets (1963), Chenery and associates (1960, 1968, 1975), Adelman and Morris (1973), Ahluwalia (1976) and Chiswick (1971). These studies share a common methodology, consisting basically of looking at a *cross-section of countries*, and (1) measuring the degree of inequality in each, (2) measuring other characteristics of each country (e.g., level of GNP, its rate of growth, importance of agriculture in total product, etc.), and (3) relating the level of inequality to that economy's characteristics using correlation or regression analysis.

In the last few years another type of approach has been followed, which looks instead at inequality *within a country*, and measures the contribution of the various components to total inequality. In this type of approach, using a variety of methodologies, inequality has been decomposed by economic sector (e.g., urban vs. rural), income source (e.g., income from labor vs. capital vs. land vs. transfers), or family characteristics (including attributes of the workers, their jobs, and regional and other locational considerations). This mode of inquiry is potentially of great value for understanding the structure of inequality and identifying which are the most important explanatory factors.

This study explores the decomposition type of inequality analysis. I summarize the alternative decomposition methodologies which have been set forth in the literature and review the principal findings of empirical studies.

II. Types of decomposition problems

Decomposition problems are of three general types: functional decomposition by income source, functional decomposition by economic sector, and microeconomic decomposition by income-determining characteristics. Let us now review each.

A. *Decomposition by functional income source*

The starting point for source decompositions is the assumption that income determination can best be studied by disaggregation by a small number of functional income sources. Take as an example the familiar functional division of income into income from labor, income from capital, and (at the micro level) income from transfers. The question asked by source decompositions is: of total inequality, how much is attributable to income from labor, how much to income from capital, and how much to income from transfers? Source decomposition procedures quantify these effects and further show how each source's contribution to overall inequality depends positively on the degree of inequality of each income source, the importance of that income source in total income, and the extent of correlation between income from that source and total income.

B. *Decomposition by economic sector*

Sectoral decompositions divide the economy into economic sectors (e.g., agriculture vs. non-agriculture). Generally, these sectors are thought to be mutually exclusive, so that all of the household's income is treated as agricultural or non-agricultural. The question asked by sector decompositions is: of total inequality, how much is attributable to variability in agricultural incomes, how much to variability in non-agricultural incomes, and how much to between-sector inequality?

Sector and source decompositions have been presented independently here, as is the practice in the literature. This distinction, though convenient, is not necessary. The economy could very easily be divided into segments defined by source-sector combinations, e.g., rural labor income, urban labor income, rural capital income, and so on.

Source and sector decompositions have in common the property that total inequality is completely accounted for by the several components, in much the same way that total national income is completely accounted for by summing income from consumption, investment, government expenditures, and net exports. The characterization of source and sector decompositions as accounting procedures is deliberate. For just as decompositions of national income into consumption, investment, government, and export components cannot explain why national income was what it was, neither can source and sector decompositions explain why national income inequality was what it was. The value of these decompositions is that they gauge the relative importance of various sources and sectors in respect to overall inequality, and thereby direct our attention to potentially fruitful areas of research.

Suppose, for instance, we find, as indeed the data show, that the primary contribution to overall income inequality is made by variation in labor income. This suggests that a valuable next step in understanding overall income inequality would be to study those economic forces which might determine the amount and distribution of labor income. In this connection, many characteristics of family members and their jobs become important. Note that microeconomic data on the individual households and their family members are needed to explore the determinants of income from labor or any other source or sector. Let us now consider what types of decompositions can be performed when such microeconomic data are available.

C. Decomposition by income determinants

A now large number of studies of less developed countries have shown that households' overall incomes and labor market earnings are systematically related to a number of family characteristics: the number of labor force participants, their incidence of unemployment, their personal characteristics (such as education and age), the family's location (by region, size of place, or rural vs. urban), the nature of their jobs (including occupation, industry, and employer's characteristics).² In a few of these studies (see Section V.C below), attempts have been made to decompose income inequality according to income determinants.

Determinant decompositions ask the question: of total inequality, how much inequality is associated with variation in income determinant 1, how much with income determinant 2, etc. and how much is not associated with any of the explanatory variables? The presence of an unexplained component is one important difference between the determinant decompositions and the other types of decompositions. Another important difference is that determinant decompositions provide much more insight into causal factors underlying the distribution of income than is the case with decompositions by source and/or sector.

We now turn to the different types of decomposition methodologies.

III. Decomposition methodologies

Four different decomposition methodologies are in current use: Gini decompositions, Theil decompositions, Analysis of Variance, and decomposition of the Atkinson index. We consider these in turn, highlighting their conceptual features.

A. Gini decomposition

The Gini coefficient is the most popular measure of relative income inequality, owing to the ease of interpreting it vis-a-vis the Lorenz curve. Gini decomposition procedures have been devised by Rao (1969), Fei and Ranis (1974, 1978), Pyatt (1976), and Mangahas (1975), among others; several of these

applications were derived independently. In addition to the empirical applications by these authors, Gini decompositions have been applied in research by Mehran (1974), Ayub (1977), and Fields (1977).

For purposes of discussion, let us suppose there are three income sources-wage income, property income, and transfer income-and that the sum of these is the total income for each family and for the economy as a whole. A decomposition by additive factor components is presented below.

Using the Gini coefficient as our measure of inequality, it might be thought that the overall Gini for the economy as a whole would be a weighted average of the factor Ginis for the individual components, the weights being given by the factor share of that income in the total. This is, however, incorrect, because the overall Gini coefficient requires the households to be ranked in increasing order of income and the different component incomes (wage, property, transfer) may not be monotonically related to one another or to the total.

To indicate the correct relationship between the overall Gini coefficient and the factor Ginis, let us order the families according to total income. For each factor income source, we may then compute a so-called pseudo-Gini coefficient, i.e., the Gini coefficient that would be obtained if households in that sector were not ordered with their incomes monotonically increasing. The overall Gini for the economy (G) turns out to be a weighted average of the pseudo-Ginis for the i th income source (\bar{G}_i) with the weights given by the factor share of that income source (ϕ_i):

$$G = \bar{G}_1\phi_1 + \bar{G}_2\phi_2 + \bar{G}_3\phi_3. \quad (1)$$

Fei, Ranis, and Kuo (1978) have shown that the pseudo-Gini for the i th source (\bar{G}_i) is equal to the product of the true Gini for that source (G_i) and a relative correlation coefficient (R_i), defined below:

$$\bar{G}_i = G_i R_i. \quad (2)$$

For each factor, the relative correlation coefficient is the ratio of two other correlations:

$$R_i = \frac{\text{cor}(Y_i, \rho)}{\text{cor}(Y_i, \rho_i)} = \frac{\text{coefficient of correlation between factor income amount and total income rank}}{\text{coefficient of correlation between factor income amount and factor income rank}} \quad (3)$$

To further explain (3), consider the R_i for labor income. The numerator of (3) is the correlation between labor income in dollars (Y_i) and the family's total income position (ρ), ordered from lowest to highest. The denominator of (3) relates the dollar labor income figure (Y_i) to that family's labor income rank (ρ_i).

Substituting (2) and (3) into (1) and dividing through by G , we obtain:

$$100\% = \phi_1 \frac{G_1}{G} \frac{\text{cor}(Y_1, \rho)}{\text{cor}(Y_1, \rho_1)} + \phi_2 \frac{G_2}{G} \frac{\text{cor}(Y_2, \rho)}{\text{cor}(Y_2, \rho_2)} + \phi_3 \frac{G_3}{G} \frac{\text{cor}(Y_3, \rho)}{\text{cor}(Y_3, \rho_3)}, = FIW_1 + FIW_2 + FIW_3, \quad (4)$$

the FIW's denoting the so-called Factor Inequality Weights of labor, property, and transfer income respectively. Overall inequality in an economy is seen to depend on the degree of inequality of each income source, the extent of correlation between income from that source and total income, and the importance of that income source in the total.

The Gini coefficient has also been decomposed in other ways. For example, Mangahas (1975) decomposed inequality into rural and urban components as follows:

$$G = \sum_j \theta_j G_j + \sum_{i>j} \phi_i \phi_j \left(\frac{D_{ij}}{\bar{Y}} \right), \quad (5)$$

where G = overall Gini coefficient

G_j = Gini coefficient among those in group j ,

ϕ_j = family income in group j as proportion of total family income,
 ϕ_j = families in group j as proportion of all families,
 \bar{Y} = mean income,
 D_{ij} = Gini difference.

The first summation measures within-sector inequality, and the second, between-sector inequality. Another breakdown was offered by Pyatt (1976) who interpreted the Gini coefficient as the expected value of a game in which a randomly-drawn individual compares his income with others'.

Other decomposition procedures partition total inequality differently. These are reviewed below.

B. Theil decomposition

A decade ago, Theil (1967) set forth a readily-decomposable inequality measure, which he subsequently (1972) illustrated with a number of empirical applications. Because an exact decomposition is possible, the Theil index has received widespread use. Among the studies of LDCs performing Theil decompositions are those by Fishlow (1972), Van Ginneken (1974), Chiswick (1976) and Uribe (1976).

The Theil index of inequality is derived rigorously from the notion of entropy in information theory. The fundamental idea of information entropy is that occurrences which differ greatly from what was expected should receive more weight than events which conform with prior expectations. The entropy index gauges the expected information content from the various outcomes, with the weights depending on the likelihood of each.

Building on this concept of entropy, the Theil index (T) of income inequality is formally the expected information of the message which transforms population shares into income shares. Mathematically, its algebraic formula is given by

$$T = \sum_{i=1}^n q_i \log \frac{q_i}{1/n}, \quad (6)$$

Where n = number of individuals or households,
 q_i = income share of i th individual.

Theil (1972, p. 100) notes that T equals the mean product of income and its own logarithm. Why this should be used as measure of economic inequality is far from transparent.

In any case, the main attraction of the Theil index lies not in its intuitive justification but rather, as remarked above, in its decomposability. Theil decompositions are well-suited for estimating the contribution of different groups to total inequality; examples of such groups are economically distinct regions of a country or population subgroups divided into educational and/or age categories.

Various decomposition formulas are given in Theil (1972, p. 100), Fishlow (1972, p. 395), Chiswick (1976, p. 9), Szal and Robinson (1977, p. 524) and Altimir and Pinera (1977, p. 14) among other places. Fishlow, for instance, gives two alternative decomposition procedures:

$$I_{ijk} = \sum_i y_{i..} \log \frac{y_{i..}}{x_{i..}} + \sum_i y_{i..} \left\{ \sum_j \frac{y_{ij.}}{y_{i..}} \log \frac{y_{ij.}/y_{i..}}{x_{ij.}/x_{i..}} \right\} + \sum_i \sum_j \frac{y_{ij.}}{y_{i..}} \left\{ \sum_k \frac{y_{ijk}}{y_{ij.}} \log \frac{y_{ijk}/y_{ij.}}{x_{ijk}/x_{ij.}} \right\} \quad (7)$$

$$I_{jk} = \sum_j y_{j.} \log \frac{y_{j.}}{x_{j.}} + \sum_k y_{.k} \log \frac{y_{.k}}{x_{.k}} + \left\{ \sum_j \sum_k y_{jk} \log \frac{y_{jk}}{x_{jk}} - \sum_j y_{j.} \log \frac{y_{j.}}{x_{j.}} - \sum_k y_{.k} \log \frac{y_{.k}}{x_{.k}} \right\}, \quad (8)$$

where y are the income shares, x the population shares, and the subscripts i , j , and k refer to income class, sector, and education. Equation (7) decomposes total inequality into between-group and within-group

components, while (8) decomposes the between-group component according to the variation among the means of the various groups.

Another decomposition procedure, substantially similar in nature, is the analysis of variance, which we now examine.

C. Analysis of variance (ANOVA)

ANOVA procedures have a long history in social scientific analysis, but their applications to economic problems are quite limited. In particular, on the problem of economic inequality in less developed countries, work is just beginning; see Langoni (1972, 1975), Chiswick (1976), Fields (1977), and Fields and Schultz (1977).

The basic idea of analysis of variance is to decompose the variance of a dependent variable, which is the sum of squared deviations from the overall mean, into two types of effects: those due to variation between different groups and those due to variation within each of the groups. For example, if the dependent variable is the logarithm of income in each of a number of households and the independent variable is the region of the country in which they live, the total sum of squares (SS) of income is expressed as:

$$SS_y = SS_{between} + SS_{within} \quad (9)$$

Where $SS_y = \sum_j \sum_i (Y_{ji} - \bar{Y})^2$

in which \bar{Y} is the overall mean of log income Y in the entire sample, the i 's are households, and the j 's are various regions

$$SS_{between\ regions} = \sum_j N_j (\bar{Y}_j - \bar{Y})^2$$

in which \bar{Y}_j is the mean log income in region j , and N_j is the number of sample households in region j

and $SS_{within\ regions} = \sum_j \sum_i (Y_{ji} - \bar{Y}_j)^2$

In other words equation (9) tells us the relative importance of income inequality within regions as compared with diversity in mean incomes across regions.

In the example of the preceding paragraph, the only explanatory factor was region. ANOVA may also handle multiple explanatory variables, say region and education. We then obtain a breakdown such as:

$$\begin{aligned} SS_y = & SS \text{ due to region} \\ & + SS \text{ due to education} \\ & + SS \text{ due to interaction between region and education} \\ & + SS \text{ within region-education groupings.} \end{aligned} \quad (10)$$

A decomposition like (10) tells us whether income inequality is greater across regions or across educational groups, whether the effects of region and education on income are independent of one another, and the relative importance of variations across these groupings are compared with the variations within them. Both gross and marginal effects may be estimated. Additionally, and quite importantly, tests of statistical significance are available for each factor.

Finally, a major advantage of analysis of variance techniques is that because they are very much like multiple regressions they indicate the quantitative importance of each category of each explanatory variable. Thus, we can learn from ANOVA decompositions how much difference it makes to one's

income if the family is located in one region rather than another or if the family head has one more year of education. No other decomposition procedure yields such information.

D. Atkinson index

The inequality index proposed by Atkinson (1970) represents the cumulative deviation of the actual income distribution from the "equally distributed equivalent income," which is the "level of income per head which if equally distributed would give the same level of social welfare as the present distribution." Without repeating the algebraic derivation, the index is given by

$$I = 1 - \frac{1}{\mu} [\sum_i f(Y_i)(Y_i)^{1-\varepsilon}]^{\frac{1}{1-\varepsilon}}, \quad (11)$$

where μ = mean income

Y_i = income of i th income recipient

ε = parameter of relative inequality aversion, as specified by the observer

It is clear with the Atkinson index (and implicit with the other measures) that value judgements are an integral part of inequality measurement. If the population is divided into mutually exclusive and jointly exhaustive groups, the Atkinson index is readily decomposable into within-group and between-group components, as follows:

$$\begin{aligned} (a) \quad I &= I_B + I_W \\ (b) \quad I_B &= 1 - \frac{1}{\mu} [\sum_j \lambda_j \mu_j^{1-\varepsilon}]^{1/(1-\varepsilon)} \end{aligned} \quad (12)$$

where λ_j = share of j th group in total population,

μ_j = mean income of j th group,

μ, ε = as before.

This kind of decomposition into within-group and between-group components closely parallels similar breakdowns of other inequality indices.

The Atkinson index is sometimes criticized as being subjective, since a value judgement (on the magnitude of the inequality aversion parameter) must be made in order to calculate the index. Yet, the Atkinson index only does explicitly what other inequality measures do implicitly: to pass judgement on the desirability of income gains at different points in the same income distribution. An index like

$$I^* = 1 - \frac{1}{\mu} [\sum_i f(Y_i)] \quad (13)$$

may *look* more objective but it is not: this is merely the Atkinson index with ε arbitrarily set equal to one. The value judgements implicit in the familiar inequality indices (like the Gini coefficient, Theil index, and log variance) are only partially understood. In recent years, the properties of these measures have been scrutinized with some care. Section IV reports on some of the relevant considerations.

IV. Choice of decomposition procedure

In weighing the advantages of the various decomposition procedures for empirical research, three central issues arise: the properties of the inequality measure itself, the richness of the information derived from the decomposition, and the suitability of the measure for the different decomposition problems.

A. *Properties of the different measures of inequality*

One way of choosing which inequality measure to decompose is to consider the measure's basic nature. In this respect, the Gini decomposition and Analysis of Variance applied to the logarithms of income come out ahead. The Gini coefficient is easily conceptualized in terms of the Lorenz curve. The variance has a familiar basis in standard statistical analysis; furthermore, income distributions are approximately log normal in shape, so analysis of the log variance is conceptually appealing. The difficulty with the Atkinson index is that it is derived from a welfare framework with which many students of inequality may disagree.⁵ Finally, the Theil index, despite its wide usage as a measure of inequality, has no clear interpretation.

Another selection criterion is the usefulness of the inequality measure in making inequality comparisons. Among the desirable axioms for this purpose are:

A1. *Axiom of scale irrelevance.* If one distribution is a scalar multiple of another (i.e., everyone's income in the first case is $x\%$ of their income in the second), then the two distributions have the same degree of inequality. Put somewhat differently, the degree of inequality in the distribution of income is measured independently of the level of income.

A2. *Axiom of symmetry.* If two income distributions are identical except that different families receive the income in the two cases, then the two distributions have the same degree of inequality. This follows from the principle of treating all individuals and families alike with regard to income distribution.

A3. *Axiom of rank -preserving equalization.* If one distribution is obtained from another by the transfer of a positive amount of income from a relatively rich family to a relatively poor one while preserving their relative rank in the distribution, then the new distribution is more equal than the old. (While few persons are likely to quarrel with this axiom, it should be noted that some additional, non-trivial assumptions about the nature of judgements of social well-being are necessary to guarantee that a "more equal" distribution is always regarded as "better.")

The Gini coefficient, Theil index, and Atkinson index satisfy these axioms. The variance does not fulfill the Axiom of Scale Irrelevance. However, Scale Irrelevance and the other axioms are satisfied by the variance of the logarithm of income (commonly known as the log-variance). Hence, all four of the inequality measures considered above are suitable by the axiomatic criterion for decomposition analysis.

Another consideration of some importance is the sensitivity of the different measures to income changes at various points in the distribution. Persons whose value judgements lead them to give greatest weight to the economic position of the poor may wish to choose that inequality measure which is most sensitive to inequality associated with low income groups. Observations on the several inequality measures may be found in Sen (1973), Weisskoff (1970), Szal and Robinson (1977), and Chiswick (1976) among others, but perhaps the most thorough analysis of this question is in the work of Champernowne (1974). He found, among other things, that the variance of the logarithms of income is most sensitive to inequality associated with poverty, the Theil index is most sensitive to inequality associated with the very rich, and the Gini coefficient is most sensitive to inequality in the middle of the income distribution. For observers whose main concern is with the low income population, analysis of variance procedures would appear more appropriate on this basis.

Let us now take up a number of other considerations which are relevant to the choice of decomposition procedure.

B. *Decomposition properties*

Here is a list of desirable output from decomposition exercises:

1. Decomposes overall inequality into within-factor and between-factor components;
2. Measures the gross contribution of each explanatory factor to total inequality;
3. Tests the statistical significance of these main effects;

4. Measures the marginal contribution of each explanatory factor;
5. Tests the statistical significance of the marginal effects;
6. Measures the effects of interactions between pairs of explanatory factors (and higher order combinations if needed);
7. Tests the statistical significance of the interaction effects;
8. Estimates the magnitude of the effect on income of each category of each explanatory variable.

ANOVA does all eight of these, Theil decompositions do only 1, 2, 4, and 6 and Gini and Atkinson decompositions only 1 and 2. Thus, in comparison with other available decomposition procedures, ANOVA provides richer information on the sources of inequality.

C. Different decomposition procedures for different decomposition problems

Consider first the problem of decomposing inequality by functional income source. As described above, procedures for using the Gini coefficient for this problem have been worked out in considerable detail. Particularly helpful is the technique for constructing Factor Inequality Weights and the breakdown of those weights into factor share, factor Gini, and correlational components (see equation (4)). In principle, ANOVA, Theil, and Atkinson procedures could be used similarly, but I have not yet seen them done in this way.

For the sectoral decomposition problem, which analyzes between- and within-sector inequality, each of the four procedures appears satisfactory. The choice among them is therefore partially dependent on the properties discussed above, and in part is a matter of convenience (depending, for example, on the availability of computer programs for the different procedures).

Finally, with respect to decompositions by income-determining factors, ANOVA and Theil techniques come out ahead. Both these procedures give a clear picture of the importance of each explanatory factor (e.g., education and region) in determining overall inequality, while at the same time gauging the unexplained residual. Gini decompositions, on the other hand, deal with deviations from predicted values in a quite cumbersome way, the difficulty being inherent in the Gini coefficient itself. The Atkinson index has not been applied to this problem as far as I know.

In the income determinant problem how do we choose between analysis of variance and Theil decompositions? I would say that two considerations work strongly in favor of ANOVA. One advantage of ANOVA is the use of log-variance as the measure of inequality. The parallel between ANOVA and multiple regressions explaining the logarithm of income permits a richer characterization of the income determination process than does Theil.⁹ We can learn, for example, by how much rural residence reduces income.¹⁰ A second overriding consideration is the availability of statistical significance tests for ANOVA but not for Theil. Thus, using ANOVA, we can measure the likelihood that the estimated contribution of an explanatory variable like region or education is a "true" effect compared with the alternative possibility that the apparent relationship is due to chance sampling. This permits us to bring the full logic of conventional statistical analysis to bear on the problem of ascertaining the determinants of inequality. From a causal (as versus an accounting) perspective, this is valuable indeed.

D. Summary

Many decomposition properties of the several inequality measures have been considered in this section. These considerations are summarized in Table 1. All in all, analysis of variance procedures based on the logarithms of income have the most desirable properties with no off-setting limitations. The use of ANOVA procedure in future research on the determinants of LDC inequality appears warranted. Concerning the choice of decomposition procedures for the types of problem under consideration, I would conclude: (1) The Gini decomposition technique is a proven method for the source problem; (2) For the sector problem, the choice of technique is a matter of some indifference, possibly, the available computer

software proving decisive, and (3) Analysis of variance dominates for decomposing inequality into the contributions of various determinantal factors.

TABLE 1
Summary of properties of four inequality measures relevant to decomposition

<i>Property</i>	<i>Inequality Measure</i>			
	<i>Gini</i>	<i>Theil</i>	<i>Log-Variance</i>	<i>Atkinson</i>
Decomposability	Yes	Yes	Yes	Yes
Intuitively appealing as a measure of inequality	Yes	No	Yes	Dubious
Axiomatically justified as a measure of inequality	Yes	Yes	Yes	Yes
Sensitivity to inequality associated with poverty	No	No	Yes	Yes
Ranks contribution of various determinants of inequality	Yes	Yes	Yes	Yes
Quantitative estimates of the magnitude of inequality determinants	No	No	Yes	No
Tests of statistical significance for estimated effects of various factors	No	No	Yes	No
Computer packages readily available	No	No	Yes	No

V. Survey of empirical findings in LDCs

The various techniques for decomposing inequality have been applied to analyses of the structure of inequality by income source, economic sector, and income-determining characteristics in a number of LDCs. Some patterns seem to be emerging from these studies. This section reviews the major results.

A. Source decompositions

The pioneering work on source decompositions in less developed countries is that of Fei and Ranis (1974) and Fei-Ranis-Kuo (1978) in their study of Taiwan. Their methodology was followed in subsequent research on Pakistan by Ayub (1977) and on Colombia by Fields (1977).

The source decompositions are based on the Gini coefficient. Taiwan's overall Gini is 0.28, which is among the lowest of all countries in the world. The source decomposition tells us which five income

TABLE 2
Decomposition of inequality in Taiwan, 1972

	<i>Wage</i>	<i>Mixed</i>	<i>Property</i>	<i>Gifts</i>	<i>Other</i>	<i>Total</i>
1 Factor Gini	0.2518	0.2968	0.4020	0.3965	0.2925	
2 Factor Share	0.582	0.275	0.093	0.046	0.004	1.000
3 Factor Inequality Weight	0.5187	0.2882	0.1322	0.0584	0.0024	1.000

Source: Fei and Ranis (1974).

sources (wage, mixed, property, gifts, and other) accounts for how much of the overall inequality. In the absence of microeconomic data, the decomposition was conducted across income groups.

The natural place to start is by looking at the Gini coefficients of the individual income sources. The reported factor Ginis are given in Row 1 of Table 2. Fei and Ranis report that property and gift income have the highest factor Ginis and therefore are least equally distributed, mixed and other are in an intermediate position, while wage income is most equally distributed. From this, we might be inclined to conclude that property and gift income account for the largest part of overall inequality and wage income the least. In actuality, these inferences would be mistaken, the reason being that we have omitted two important factors from consideration, namely, (1) the factor shares, which tell us the importance of that factor in total income, and (2) the correlations between factor income and total income, which tell us the degree to which that factor augments total inequality or offsets the inequality attributable to other sources.

When one looks at the factor shares in Row 2 of the Table, wage income is seen to be the most important source of income by far, mixed income is in an intermediate position, and property and gift income are relatively unimportant. As the decomposition procedure (equation (4)) showed, total inequality is a weighted average of inequality in the individual factor incomes. In the case of Taiwan, wage income is relatively equally distributed but has the largest factor share, property and gift income are relatively unequally distributed but have small factor shares, and mixed and other sources are in the middle in both respects.

The Factor Inequality Weights presented in Row 3 measure each factor's contribution to total inequality. The data show that wage income is the source of more than half of total inequality, while property and gifts combined account for less than 20%; the rest is accounted for by mixed income, some substantial but unknown part of which reflects returns to labor.

The same basic decomposition methodology has been applied to the cases of Pakistan (Ayub, 1977) and urban Colombia (Fields, 1977) with quite similar qualitative results.¹⁷ Both authors found: (1) The highest factor Gini coefficients for non-labor income sources than for labor incomes; (2) The reverse ordering for factor income shares; and (3) The overwhelming importance of labor incomes (including wage employment and self-employment) in accounting for overall inequality.

Individually and together, the results for Taiwan, Pakistan, and Colombia give a common impression about the contribution of the various income sources to overall inequality: the bulk of income inequality is attributable to labor income. The high Factor Inequality Weights for labor incomes suggest that the principal inequality-producing factor is that some people receive a great deal more income for their work than do others. This has important implications both for research (study the labor market) and for policy (create more well-paying jobs). The intuitive prior notion that the most unequally-distributed factors (property, gifts, etc.) contribute the most to total inequality is found to be false in each case.

B. Sector decompositions

Sector decomposition studies do three things: they measure the inequality within each sector or region of an economy, indicate the importance of within-sector inequality for all sectors taken together, and determine the amount of inequality accounted for by between-sector variation. The available studies decompose inequality within a country and within regions of the world.

Within-country sector decompositions have been carried out using the Gini coefficient by Mehran (1974) for Iranian cities, by Mangahas (1975) for areas and regions of the Philippines, and by Pyatt (1976) for urban and rural locations in Sri Lanka. In other studies-by Fishlow (1972, 1973) and Langoni (1972, 1975) in Brazil, van Ginneken (1975) in Mexico, Chiswick (1976) in Thailand, Fields (1977) and Fields and Schultz (1977) in Colombia, and Altimir and Pinera (1977) in Chile, Panama, and Venezuela-regional or urban-rural decompositions were undertaken as part of a larger exercise; these studies used Theil decompositions or analysis of variance. Without exception, the result emerges that variations *within* sectors or regions are far more important in accounting for inequality than variations *between* sectors.

Another result of the within-country sector decompositions is that inequality is found to be greater within urban than within rural areas. See, for example, Mangahas (1975, p. 295) for the Philippines, Pyatt (1976, Table 3) for Sri Lanka, Fei-Ranis-Kuo (1978, Diagram 2) for Taiwan, Ayub (1977, Table XII) for Pakistan, and Fields and Schultz (1977, Table 4) for Colombia. These results accord with the findings of Kuznet (1955) and many other income inequality studies.

Sector decompositions have also been applied to studies of inequality in the world. First Theil (1972) and after him Uribe (1976) using the same methodology examined the structure of inequality within a number of countries and across countries. Theil's analysis covered all parts of the world, while Uribe's was limited to Latin America only. Both studies found more intra- than intercountry inequality.

In summary, the sector decomposition studies report more inequality *within* sectors or countries than *across* them. As with the source decomposition literature, these studies clearly demonstrate the importance of going down to the household level in order to understand the determinants of incomes and income inequality.

C. Determinant decompositions

Eight studies decomposing inequality in less developed countries by income determinants are surveyed here. The countries covered are Brazil (two studies), Mexico (two studies), Thailand, Taiwan, Colombia (three studies), Argentina, Costa Rica, Chile, Panama, Peru, and Venezuela. Three of the four statistical decomposition methodologies (excluding the Atkinson index) have been used. The results of these studies are summarized in Table 3.

The available studies exhibit several similarities: (1) Large effects are found for personal attributes. (2) Of the personal attributes considered, education and age contribute roughly equal explanatory power. (3) Large effects are sometimes found for employment aspects. (4) Of the employment aspects considered, the most important correlate of income is occupation. (5) Regional effects are found to be of some importance, but these effects are not major ones. (6) Intra-regional inequality dominates inter- regional inequality.

The considerable importance of personal attributes in the decomposition studies accords with the findings of income- and earnings-generating functions; see, for example, Fields (1978) and the reference cited therein for Colombia; McCabe (1974) for Zaire; Langoni (1975) for Brazil; Johnson (1971) for Kenya; and Chiswick (1976) for Thailand. In those studies, personal characteristics were found to explain as much as 60% of the variance in the logarithms of income. Little was gained by adding information on the employer or the place of residence.

The decomposition findings of Altimir and Pinera differ from the others in attributing a large role to employment variables, particularly occupation. This contrasts with earnings function studies which typically find a small effect of occupation or which have omitted occupation entirely on the grounds of its

presumed endogeneity. The Altimir-Pifera results on this point are suspect because of the decomposition equivalent of simultaneous equations bias. It seems warranted to conclude that occupation and other employment variables help explain income inequality but that these variables have lesser effects than do personal variables.

Other sources also suggest the limitations of analyses of income distribution at the sectoral level. Webb (1976), for instance, reports that the poor in Lima are found scattered in many different sectors-commerce, manufacturing, transport, construction, public service, modern sector firms or occupations, and miscellaneous services-each sector containing at least 10% of the poor. More generally, it would appear that to predict an individual's income, we can do much better knowing his education and age than which economic sector he is located in.

Decompositions of inequality by income-determining characteristics, such as those summarized in Table 3, are potentially of great usefulness in analyzing LDC wage structures. Economic theory does not yet offer a comprehensive explanation for income inequality. However, we do have partial explanations based on considerations of labor demand, labor supply, technological variability, and institutional influences. Attempts to integrate these various strands of analysis into a unified theory of the determinants of wages and size distribution of income and to implement such a theory empirically have met with only partial success. The empirical results of decomposition studies may aid in the inductive development of a more comprehensive view of this vitally important process.

VI. Conclusions

This paper has considered three types of decompositions of inequality and four methodologies for decomposition analysis and reviewed the findings from empirical studies in less developed countries. Several methodological and empirical conclusions emerge:

1. *The three different decompositions (by functional income source, by economic sectors, and by income-determining characteristics) are basically quite different.* The first two types of decompositions give a total accounting for inequality, whereas determinant decompositions allow for an unexplained residual component. Also, source and sector decompositions are more of an accounting nature, while determinant decompositions are interpreted causally. Finally, an important difference between source decompositions and sector decompositions is that many households receive income from more than one source, but not ordinarily from more than one sector.
2. *The various decomposition methodologies (by Gini coefficient, Theil index, analysis of variance and Atkinson index, are suited for different types of problems.)* For the source problem, the Gini decomposition technique is an effective method; I have not seen the decompositions of the other measures applied in this way. In analysis of inequality within and among mutually exclusive sectors, any of the available techniques will serve satisfactorily, although if tests of statistical significance are of interest, analysis of variance may be preferable. For gauging the causal importance of various explanatory factors, analysis of variance can do more than any of the other approaches. ANOVA may also be preferred for its greater sensitivity to income inequality associated with the poverty population.
3. *Source decomposition studies point to variation in labor incomes as the predominant factor accounting for income inequality.* To understand the structure of income inequality in LDCs, knowledge of the determinants of income from wages and self-employment becomes paramount, as does an understanding of the functioning of LDC labor markets.
4. *Sector decomposition studies indicate substantially more inequality within regions than across them.* This implies the need to look within regions for other sources of income variability, at the level of either the worker or his job. Empirically, simple models of dualistic economic development and labor market segmentation will not do.

5. *From studies which decompose inequality by income-determining characteristics, we find that more inequality is attributable to variation in personal characteristics than to the sector of employment or locational aspects.* The most powerful personal characteristics explaining inequality are education and age. Occupation, economic sector, and location make some contribution to explaining inequality, but these variables have lesser effects.
6. *Singly and together, decomposition studies in less developed countries lead to an inescapable conclusion: the overwhelming importance of income variation according to attributes of individuals and the secondary role of variation between economic segments grouped according to sector of the economy or functional income source.* Given this overall conclusion, the need for further microeconomic income determination studies at the level of the household stands out. Sectoral considerations may have a role to play in determining LDC inequality too, explaining why some individuals with a given set of personal attributes (education, age, sex, etc.) receive higher incomes than others. These studies, when combined with more macroeconomic analyses, may shed some light on the systemic forces generating inequality in LDCs.

TABLE 3
Decomposition of inequality in ten less developed countries by income determinants, rank ordering of findings

Country, date, author, and decomposition methodology		Factors considered, in order of importance			
BRAZIL, 1960					
Fishlow (1972)	1. Education				
Theil Decomposition	2. Age				
	3. Sector				
	4. Region				
BRAZIL, 1960 and 1970					
Langoni (1975)		1960	1970		
Multiple Regression Approach to Analysis of Variance (ANOVA)		Rank	Rank		
	Education	1	1		
	Region	2	3†		
	Age	3	2		
	Sex	4	3†		
	Activity	5	3†		
THAILAND, 1971 (URBAN)					
Chiswick (1976)	Theil, Gross Explanation	Theil, Marginal Explanation	ANOVA, Gross Explanation	ANOVA, Marginal Explanation	
Theil Decomposition and Analysis of Variance (ANOVA)					
	Age	1	2	2**	3**
	Education	2	1	1**	1**
	Sex	3	4	3*	2**
	Farm Occup.	4	5	4x	4**
	Type of Earnings	5	3	5x	5**
TAIWAN, 1966					
Fei, Ranis, and Kuo (1978)	1. Education				
Gini Decomposition	2. Age				
	3. Sex				
COLOMBIA, 1967/68 (URBAN)					
Fields (1977)	1. Education				
Analysis of Variance (ANOVA)	2. Age				
	3. City				
COLOMBIA, 1973					
Fields and Schultz (1977)	1. Education**				
Analysis of Variance (ANOVA)	2. Age**				
	3. Region**				
	4. Urban/Rural**				
	5. Type of Employment**				
COLOMBIA, 1974 (URBAN)					
Altimir and Piñera (1977)		Gross Explanation	Marginal Explanation		
Theil Decomposition					
	Education	1	1		
	Occupation	2	4†		
	Age	3†	4†		
	Kind of econ. activity	3†	2		
	Employment status	5	3		
	Sex	6	7		
	Time Worked	7	6		

TABLE 3 (Contd.)

Country, date, author, and decomposition methodology		Factors considered, in order of importance	
MEXICO, 1968			
van Ginneken (1975)		1. Education	
Theil Decomposition		2. Urban-Rural	
		3. Age	
		4. Sector of Activity	
		5. Occupation	
MEXICO, 1970			
Altimir and Piñera (1977)		Gross	Marginal
Theil Decomposition		Explanation	Explanation
Education		1	2
Occupation		2	3†
Age		3	1
Sex		4†	6
Employment status		4†	5
Kind of econ. activity		6	3†
COSTA RICA, 1966/67			
Altimir and Piñera (1977)		Gross	Gross
Theil Decomposition		Explanation	Explanation
Education		1†	2
Occupation		1†	3
Age		3	1
Sex		4†	4†
Employment status		4†	6†
Kind of econ. activity		6	6†
Time worked		7	4†
CHILE, 1971			
Altimir and Piñera (1977)		Gross	Marginal
Theil Decomposition		Explanation	Explanation
Occupation		1	1
Age		2	2
Employment status		3	4†
Sex		4	4†
Kind of econ. activity		5	3
Time worked		6	6
Education		N.A.	N.A.
PANAMA, 1972			
Altimir and Piñera (1977)		Gross	Marginal
Theil Decomposition		Explanation	Explanation
Education		1	1†
Occupation		2	3†
Time worked		3	3†
Age		4	1†
Kind of econ. activity		5	3†
Employment status		6†	6†
Sex		6†	6†

TABLE 3 (Contd.)

Country, date, author, and decomposition methodology		Factors considered, in order of importance	
<i>PERU, 1970</i>			
Altimir and Piñera (1977) Theil Decomposition		Gross Explanation	Marginal Explanation
	Education	1	2
	Occupation	2	4†
	Age	3	1
	Kind of econ. activity	4†	4†
	Employment status	4†	6
	Sex	4†	7
	Time worked	7	3
<i>VENEZUELA, 1971</i>			
Altimir and Piñera (1977) Theil Decomposition		Gross Explanation	Marginal Explanation
	Education	1	2
	Occupation	2	3
	Age	3	1
	Employment status	4	4
	Sex	5	7
	Time worked	6	5
	Kind of econ. activity	7	6

Notes to Table 3

** = Statistically significant effect at 0.01 level

* = Statistically significant effect at 0.05 level

x = Not statistically significant effect at 0.05 level

If no **, *, or x appears, no test of statistical significance is possible

† Contributions of these variables were virtually identical

BIBLIOGRAPHY

- ADELMAN, I. and MORRIS, C. T., *Economic Growth and Social Equity in Developing Countries*, Stanford, 1971.
- AHLUWALIA, M. S., "Inequality, Poverty and Development," *Journal of Development Economics*, December, 1976.
- ALTIMIR, O. And PINERA, S., "Decomposition Analysis of the Inequality of Earnings in Latin American Countries," Economic Commission for Latin America and World Bank, mimeo, August, 1977.
- ATKINSON, A. B., "On the Measurement of Inequality," *Journal of Economic Theory*, Vol. 2, 1970.
- AYUB, M., *Income Inequality in a Growth-Theoretic Context: The Case of Pakistan*, Unpublished doctoral dissertation, Yale University, 1977.
- CHAMPERNOWNE, D. G., "A Comparison of Measures of Inequality of Income Distribution," *The Economic Journal*, December, 1974.
- CHENERY, H. B., "Patterns of Industrial Growth," *American Economic Review*, September, 1960.
- and TAYLOR, L., "Development Patterns: Among Countries and Over Time," *Review of Economics and Statistics*, November, 1968.
- and SYRQUIN, M., *Patterns of Development, 1950-1970*, London, 1975.
- CHISWICK, B., "Earnings Inequality and Economic Development," *Quarterly Journal of Economics*, February, 1971.
- CHISWICK, C. U., "Application of the Theil Index to Income Inequality," LB.RD. Working Paper Series B-2, July, 1976.
- DAS, T. and PARIKH, A., "Decomposition of Inequality Measures and a Comparative Analysis," School of Social Studies, University of East Anglia, Discussion Paper No. 44, May, 1977.
- FEI, J. C. H. and RANIS, G., "Income Inequality by Additive Factor Components," Economic Growth Center, Yale University, Center Discussion Paper No. 207, June, 1974.
- , ----- and Kuo, S. W. Y., "Growth and the Family Distribution of Income by Factor Components," *Quarterly Journal of Economics*, February, 1978.
- FIELDS, G. S., "Income Inequality in Urban Colombia: A Decomposition Analysis," Economic Growth Center, Yale University, Discussion Paper No. 267, September, 1977.
- , "Analyzing Colombian Wage Structure," World Bank Studies in Employment and Rural Development No. 46, May, 1978.
- and FEI, J. C. H., "On Inequality Comparisons," *Econometrica*, March, 1978.
- and SCHULTZ, T. P., "Sources of Income Variation in Colombia: Personal and Regional Effects," Economic Growth Center, Yale University, Center Discussion Paper No. 262, June, 1977.
- FISHLOW, A., "Brazilian Size Distribution of Income," *American Economic Review*, Vol. 62, No. 2, May, 1972.
- , "Distribuicao da Renda no Brasil: Um Novo Exame," *Dados*, No. 11, 1973.
- JOHNSON, G., "The Determination of Individual Hourly Earnings in Urban Kenya," Institute for Development Studies, University of Nairobi, Discussion Paper No. 115, September, 1971.
- KUZNETS, S., "Economic Growth and Income Inequality," *American Economic Review*, Vol. 45, No. 2, March, 1955.
- , "Quantitative Aspects of the Economic Growth of Nations: VIII. Distribution of Income by Size," *Economic Development and Cultural Change*, Part II, Vol. 11, No. 2, January, 1963.
- LANGONI, C., "Distribuicao da Renda e Desenvolvimento Economico do Brasil," *Estudos Economicos*, Outubro, 1972.
- , "Income distribution and Economic Development: The Brazilian Case," Paper presented at the World Econometric Society Congress, Toronto, 1975.
- MCCABE, J., "The Distribution of Labor Incomes in Urban Zaire," *Review of Income and Wealth*, Vol. 20, March, 1974.

- MANGAHAS, M., "Income Inequality in the Philippines: A Decomposition Analysis," in *Income Distribution, Employment, and Economic Development in Southeast and East Asia*, Papers and Proceedings of the seminar sponsored jointly by the Japan Economic Research Center and the Council for Asian Manpower Studies, July, 1975.
- MEHRAN, F., "Decomposition of the Gini Index: A Statistical Analysis of Income Inequality," International Labour Office, mimeo, 1974.
- PAUKERT, F., "Income Distribution at Different Levels of Development: A Survey of Evidence," *International Labour Review*, August-September, 1973.
- PYATT, G., "On the Interpretation and Disaggregation of Gini Coefficients," *The Economic Journal*, Vol. 86, June, 1976.
- RAO, V. M., "Two Decompositions of Concentration Ratios," *Journal of the Royal Statistical Society*, 1969.
- SEN, A. K., *On Economic Inequality*, Norton, 1973.
- SZAL, R. J. and ROBINSON, S., "Measuring Income Inequality," in C. R. Frank and R. C. Webb, eds., *Income Distribution and Growth in the Less-Developed Countries*, Brookings Institution, 1977.
- THEIL, H., *Economics and Information Theory*, Amsterdam: North Holland Press, 1967.